

CGCGCTGCGCGGACGGGACGCCAGCAGCGGAGGCGCGCGCCAGCAGACCCCGGGACCATG	50
M	1
GGCTCCATGTTCTGGAGCGGAGGAGGTGGCGCTGGTCCAGCTCTTTCTGCCACAGCGGT	120
G S N F R S E E V A L V Q L F L P T A A	21
GCCTACACCTGCGTGAGTGGCGTGGCGGAGCTGGCGCTCGTGGAGTTCAGAGACCTCAAC	180
A Y T C V S R L G E L G L V E F R D L N	41
GCCTCGGTGAGCGCTTCAGAGAGCGCTTGTGGTTCATGTTTGGCGCTGTGAGGAGCTG	240
A S V S A F Q R R F V V D V W R C E E L	61
GAGAAGACCTTCACCTTCTGCAAGGAGGAGGTGGCGCGGTGGCGTGGTCTGCCCCCG	300
E K T F T F L Q E E V R R A G L V L P P	31
CCAAAGGGGAGGCTGCGCGCACCCCGACCCCGGACCTGCTGCGCATCCAGGAGGAGACG	360
P K G R L P A P P P R D L L R I Q E E T	101
GAGCGCTGCGCCAGGAGCTGCGGATGTGCGGGCAACCAGCAGCGCTTGGCGGCGCAG	420
E R L A Q E L R D V R G N Q Q A L R A Q	121
CTGCACCGAGTGCAGCTCCACGCGCGCGCTGCTACGCCAGGGCCATGAACCTCAGCTGGCA	480
L H Q L Q L H A A V L R Q G H E P Q L A	141
GCCGCCCCACAGATGGGGCTCAGAGAGGAGCGCGCTGCTCCAGGCGCGCGGGGGCGG	540
A A H T D G A S E R T P L L Q A P G G P	161
CACCAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCCACAAGGCGCGTGGC	600
H Q D L R V N F V A G A V E P H K A P A	181
CTAGAGCGCTGCTCTGGAGGGCTGCGCGCGCTTCTCATTTGCCAGCTTCAGGGAGCTG	560
L E R L L W R A C R G F L I A S F R E L	201
GAGCAGCGCTGGAGCACCCCGTGACGGCGGAGCCAGCCAGTGGATGACCTTCTTCATC	720
E Q P L E H P V T G E P A T W M T F L I	221
TCCTACTGGGGTGAGCAGATCGGACAGAAGATCCGCAAGATCAGGACTGCTTCCACTGC	780
S Y W G E Q I G Q K I R K I T D C F H C	241
CAGCTCTTCCCGTTTCTGCAGCAGGAGGAGGCGCGCTGCGGGCGCTGCAGCAGCTGCAA	840
H V F P F L Q Q E E A R L G A L Q Q L Q	261
CAGCAGAGCCAGGAGCTGCAGGAGGTCTCGGGGAGACAGAGCGGTTCCTGAGCCAGGTG	900
Q Q S Q E L Q E V L G E T E R F L S Q V	281
CTAGGCGGGTGCTGCAGCTGCTGCGCGCAGGGCAGGTGCAGGTCCACAAGATGAAGGCC	960
L G R V L Q L L P P G Q V Q V H K M K A	301
GTGTACCTGGCGCTGAACCAAGTGCAGCGTGAGCACCACGCACAAGTGCTCATTTGCCGAG	1020
V Y L A L N Q C S V S T T H K C L I A E	321
GCCTGGTGTCTGTGCGAGACCTGCGCGCGCTGCAAGAGGCGCTGCGGACAGCTCGATG	1080
A W C S V R D L P A L Q E A L R D S S M	341
GAGGAGGGAGTGAGTGGCTGCGCTCACCGCATCCCGTGGCGGACATGCCCCCACACTC	1140
E E G V S A V A H R I P C R D M P P T L	361
ATCCGCACCAACCGCTTCACGCGCAGCTTCCAGGGCATCGTGGATCGCTACGGCGTGGGC	1200
I R T N R F T A S F Q G I V D R Y G V G	381
CGCTACCAGGAGGTCAACCCCGCTCCCTACACCATCATCCTTCCCGTTCTGTCTGCT	1260
R Y Q E V N P A P Y T I I T F P F L F A	401

FIGURE 1A

GTGATGTTCCGGGATGTGGGCCACGGGCTGCTCATGTTCTCTTCGGCCTGGCCATGGTC	1320
V M F G D V G H G L L M F L F A L A M V	421
CTTCCGGAGAACCGACCGGCTGTGAAGCCCGCAGAACGAGATCTGGCAGACTTCTTC	1380
L A E N R P A V K A A Q N E I W Q T F F	441
AGGGGCGGCTACCTGCTCTCTTATGGGCTGTCTTCATCTACACCGGCTTCATCTAC	1440
R G R Y L L L L M G L F S I Y T G F I Y	461
AACGAGTGCTTCAGTCGCCACACGATCTTCCCCCTCGGGCTGGAGTGTGGCCGCCATG	1500
N E C F S R A T S I F P S G W S V A A M	481
GCCAACCACTCTGGCTGGAGTGATGCATTCCTGGCCCCAGCACAGATGCTTACCTGGAT	1560
A N Q S G W S D A F L A Q H T M L T L D	501
CCCAACGTACCGGCTGTCTCTCTGGGACCTACCCCTTTGGCATCGATCTTATTGGAGC	1620
P N V T G V F L G P Y P F G I D P I W S	521
CTGGCTGCCAACCACTTGAOCTTCTCAACTCTTCAAGATGAAGATGTCGGTCATCCTG	1680
L A A N H L S F L N S F K M K M S V I L	541
GGCGTGCTGCACATGGCCTTTGGGGTGGTCTCGGAGTCTTCAACCACGTGCACCTTGGC	1740
G V V H M A F V L G V F N H V H F G	561
CAGAGGCACCGGCTGCTGCTGGAGACGCTGCCGGAGCTCACCTTCTGCTGGGACTCTTC	1800
Q R H R L L L L E T L P E L T F L L G L F	581
GGTTACCTCGTGTTCTTAGTCACTACAAGTGGCTGTGTGCTGGGCTGCCAGGGCGGCC	1860
G Y L V F L V I Y K W L C V W A A R A A	601
TGCCCCAGCATCTCTACCTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAAC	1920
S P S I L I H F I N M F L F S H S P S N	621
AGGCTGCTCTACCCCCCGCAGGAGGTGGTCCAGGCCACGCTGGTGGTCTGCGCTTGGCC	1980
R L L Y P R Q E V V Q A T L V V L A L A	641
ATGGTGGCCATCTGCTGCTTGGCACACCCCTGCACTGCTGCACCGCCACCGCCCGCCG	2040
M V P I L L L G T P L H L L H R H R R R	661
CTGGCGAGGAGGCCCCCTGACCGACAGGAGGAAAACAAGGCCGGGTTGCTGGACCTGCT	2100
L R R R P A D R Q E E N K A G L L D L P	681
GACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAGGCAGGGGGCTGGATGATGAA	2160
D A S V N G W S S D E E K A G G L D D E	701
GAGGAGGCCGAGCTCGTCCCCCTCCGAGGTGCTCATGCCACAGGCCATCCACACCATGGAG	2220
E E A E L V P S E V L M H Q A I H T I E	721
TTCCTGCTGGGCTGCGTCTCCAACACCGGCTCTACCTGCGGCTGTGGGCGCTGAGCCTG	2280
F C L G C V S N T A S Y L R L W A L S L	741
GCCCCGCCCCAGCTGTCCGAGGTTCTGTGGGCCATGGTGATGCCCATAGGCTGGGCGCTG	2340
A H A Q L S E V L W A M V M R I G L G L	761
GGCCGGGAGGTGGGCGTGGGCGCTGTGGTCTGCTCCCATCTTTGCCGCTTTGCCGCTG	2400
G R E V G V A A V V L V P I F A A F A V	781
ATGACCGTGGCTATCTGCTGCTGATGGAGGACTCTCAGCCTTCTGACGCCCTGGG	2460
M T V A I L L V M E G L S A F L H A L R	801
CTGCACTGGGTGGAATTCAGAACAAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTCCC	2520
L H W V E F Q N K F Y S G T G Y K L S P	821
TTCACTTCCCTGCCACAGATGACTAGGGCCCCACTGCAGGTCTGCCAGACCTCTCTCT	2580
F	841
GACCTCTGAGGCAGGAGAGGAATAAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAAAAAA	2640

FIGURE 1B

MG - LFRSEE - - L - QLFL - - - AAY - CVS - L - ELG - V - FRDLN - - V - - FQR -  
 FV - EVRRCEMD - - L - F - - - EIR - A - - - - - - - - - P - PROM - - - - -  
 - E - - - - EL - EI - - NQ - AL - - - F - L - - - - - ILR - - - - - A - D - - - E  
 - - - LL - - - - G - - - - LR - - - FVAG - I - - - - P - FERMLWR - CRG - - - - -  
 E - E - PLE - PVTGD - - - - - F - I - F - GDQ - - - - V - KI - E - F - - - YP - - -  
 - - - R - - - - - - - - - DLQ - VL - - - TE - - - - - VL - - - - - I - V - KM  
 KAIY - - - LN - C - I - - T - KCLIAE - WC - V - DL - - - - Q - ALR - - - - G - - V - -  
 I - - R - - - - - PPT - - - - TN - FT - - - FQ - IVD YGIG - Y - EINPAPYTIITFP  
 FLFAVMFGD - GHG - LM - LEA - - - MVL - E - R - - - - - NE - F - - - F - GRY - -  
 LLMG - FSIYTGLIYNDGFS - - - - IF - S - WSV - - - M - - - - - W - E - - L - - - -  
 L - L - P - V - GVF - GPYPFGIDPIW - - A - N - L - L - FLNSFKMKMSVILGIIHM -  
 FGV - L - - - - FNH - - - - - F - - - - - PE - - - FM - - - LFGYLV - LI - YKW - - - Y -  
 A - - - - - PS - LIHFINNFLS - - - S - N - MLY - - Q - - IQ - - LVV - AL - - VP  
 - MLL - - PL - L - - - - R - - - - - - - - - R - - - - - F - - A - - - - -  
 - - S - - - E - A - - - - EDE - - - DF - - - D - M - HQAIHTIEYCLGCISNTASYLRL  
 WALSLAHAQLSEVLW - MV - - IGL - - - - - G - - - - - F - - IFAAFA - LTVAI  
 LLIMEGLSAFLHALRLHWVEFQNKFY - GTGFKF - PP - - - - - - - - - - -

FIGURE 2